

<https://helda.helsinki.fi>

---

## Accelerated pharmaceutical protein development with integrated cell free expression, purification, and bioconjugation

Richardson, Dominique

2018-08-10

---

Richardson , D , Itkonen , J , Nievas , J , Urtti , A & Casteleijn , M G 2018 , ' Accelerated pharmaceutical protein development with integrated cell free expression, purification, and bioconjugation ' , Scientific Reports , vol. 8 , no. 1 , 11967 . <https://doi.org/10.1038/s41598-018-30435-4>

---

<http://hdl.handle.net/10138/239090>

<https://doi.org/10.1038/s41598-018-30435-4>

---

cc\_by

publishedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

# SCIENTIFIC REPORTS

OPEN

## Accelerated pharmaceutical protein development with integrated cell free expression, purification, and bioconjugation

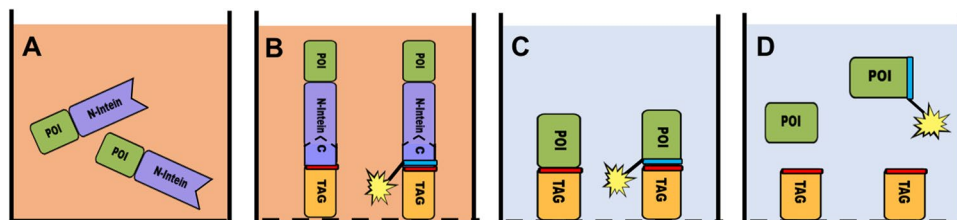
Dominique Richardson<sup>1</sup>, Jaakko Itkonen<sup>1</sup>, Julia Nieves<sup>1,2</sup>, Arto Urtti<sup>1,3,4</sup> & Marco G. Casteleijn<sup>1</sup> 

The use of living cells for the synthesis of pharmaceutical proteins, though state-of-the-art, is hindered by its lengthy process comprising of many steps that may affect the protein's stability and activity. We aimed to integrate protein expression, purification, and bioconjugation in small volumes coupled with cell free protein synthesis for the target protein, ciliary neurotrophic factor. Split-intein mediated capture by use of capture peptides onto a solid surface was efficient at 89–93%. Proof-of-principle of light triggered release was compared to affinity chromatography (His<sub>6</sub> fusion tag coupled with Ni-NTA). The latter was more efficient, but more time consuming. Light triggered release was clearly demonstrated. Moreover, we transferred biotin from the capture peptide to the target protein without further purification steps. Finally, the target protein was released in a buffer-volume and composition of our choice, omitting the need for protein concentration or changing the buffer. Split-intein mediated capture, protein *trans* splicing followed by light triggered release, and bioconjugation for proteins synthesized in cell free systems might be performed in an integrated workflow resulting in the fast production of the target protein.

In traditional biotechnology, cells are utilized for a variety of different processes<sup>1</sup>, and more recently during the eras of genetics and molecular biology, one important process is recombinant protein production<sup>2–5</sup>. For pharmaceutical protein production and screening the use of cells is a common practice and the state-of-the-art<sup>6–8</sup>. The bioprocess for recombinant protein production is divided in an upstream and a downstream process: cells are fermented before and during protein production (upstream) and the product is either secreted from the cells into the cultivation media, or cells are lysed and the cell lysates processed further (downstream). Downstream processes may include: separation of soluble and insoluble cell debris and media components, protein purification, protein formulation (including concentrating), bioconjugation, protein-bioconjugate purification, and protein refolding. Depending on the protein and the needs either all, some, or just one of these aforementioned downstream processes are required<sup>9</sup>. Proteins are prone to degradation due to many triggers, such as pH, salt concentration, organic solvents, shearing, interaction with surfaces and interfaces (including protein aggregates), lyophilisation, humidity levels, protein concentration, and temperature changes. Therefore, these processes must be designed carefully to produce intact and functional proteins.

Regarding the physicochemical basis for protein purification we can identify low, medium and high-resolution methods<sup>10</sup>. Low resolution methods are based on precipitation due to differences in solubility, for example in different ammonium sulphate, polyethylene glycol or polyethyleneimine concentrations<sup>11</sup>, based on their pI, or in combination with affinity precipitation. Phase partitioning is based on the solubility of the proteins in regards to different liquid phases (e.g. aqueous two- or three-phase partition), but the resolution is still low to medium. These processes are very useful as the first steps of a protein purification method evaluation or if the final product is of low monetary value since high resolution methods are relatively more expensive. However, usually

<sup>1</sup>Drug Research Program, Division of Pharmaceutical Biosciences, Faculty of Pharmacy, University of Helsinki, Helsinki, Finland. <sup>2</sup>Division of Pharmaceutical Chemistry and Technology, Faculty of Pharmacy, University of Helsinki, Helsinki, Finland. <sup>3</sup>School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland. <sup>4</sup>Institute of Chemistry, St Petersburg State University, Petergoff, St Petersburg, Russian Federation. Dominique Richardson and Jaakko Itkonen contributed equally to this work. Correspondence and requests for materials should be addressed to M.G.C. (email: [marco.casteleijn@vti.fi](mailto:marco.casteleijn@vti.fi))



**Figure 1.** Schematic representation of the workflow. (A) Cell free protein synthesis (CFPS) of the protein of interest (POI) fused with an N-terminal split intein, typically 4–6 hours. The CFPS matrix is depicted in beige; (B) Split intein mediated capture, for 3 hours. The capture peptide is either in solution or immobilized to the surface (the dashed line depicts both scenarios). In orange the ‘Tag’ can be any affinity tag. The red bar is a photo-cleavable amino acid, while the blue bar is an unnatural amino acid linked to a reactive moiety. At this stage the intein will splice itself out spontaneously, without co-factors<sup>46,47</sup>; (C) After the intein reaction the CFPS matrix and side-products are washed away in under half an hour, and the final buffer is added (light blue); (D) after light triggered release (3 hours) POI is released into its final formulation. Bioconjugation can be performed now, later (for example *in-vivo*) or also at the end of step C before light treatment.

pharmaceutical proteins are purified utilizing high resolution methods based on chromatographic separations. Proteins are either separated based on charge (ion exchange), hydrophobicity (hydrophobic interaction chromatography (HIC)<sup>12</sup> or reverse-phase high performance liquid chromatography (RP-HPLC), size exclusion chromatography (also called gel filtration), and binding (purification with a protein-tag or utilizing a binding region of the native protein)<sup>13,14</sup>. The current trend in protein production is to increase protein yields at every stage and to reduce the number of separation steps<sup>10</sup>.

Cell free protein synthesis (CFPS) has shown its importance for pharmaceutical protein production, including the industrial production of a commercial protein drug (rhGM-CSF)<sup>15</sup>. However, this product was still purified using traditional methods, such as weak anion exchange chromatography, followed by filtration and several size exclusion chromatography steps, with all their problems as stated before. Furthermore, rapid methods are needed for pharmaceutical protein synthesis and screening<sup>16</sup>, but so far simple and integrated workflows for protein synthesis, purification, bioconjugation, and formulation do not exist. Since each protein is unique, such a workflow should focus on the intended use of the protein<sup>5</sup>.

In this contribution, we demonstrate a workflow for the synthesis, purification, bioconjugation, and formulation of a pharmaceutical protein: human ciliary neurotrophic factor (hCNTF). Our approach utilizes cell free protein synthesis, split-intein mediated capture, light triggered release, moiety transfer upon release to facilitate bioconjugation, and formulating protein in the correct buffer at the desired protein concentration (see Fig. 1). This workflow omits many time-consuming steps of downstream protein production (e.g. the use of living cells, chromatography, dilution and re-concentrating, buffer exchange, purification of the bioconjugates, and re-folding). In principle, the proposed fast workflow can be performed in one day, instead of weeks that are spent using traditional methods. The proof-of-principle for this method is demonstrated here using ciliary neurotrophic factor (CNTF) and green fluorescent protein (GFP).

## Results

The following results are presented according to the workflow depicted in Fig. 1. However, several of the steps were also evaluated in together, in an integrated way.

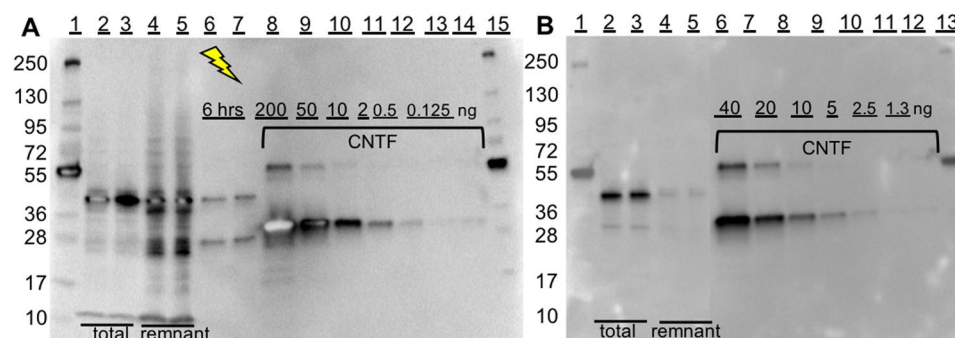
**Cell free protein expression.** Several cell-free systems were evaluated: *E. coli* BL21(DE3) S12 and S30 lysates (Figs S1 and S2), tobacco plant BY-2 cell lysates (BYL)<sup>17</sup>, commercial kits based on *E. coli* (Promega, USA; Figs S1 and S2), wheat germ (Biotech rabbit, Germany; Fig. S3), and mammalian HeLa cells (ThermoFisher scientific, USA). BYL and HeLa systems (both batch processes) gave the highest yields within 6 hours and were used for further evaluation.

Protein yields were determined from the western blots (WB; Fig. 2A,B) and are summarized in Table 1. Due to the limitations in the linear range of the WB-staining reagent the bands with highest concentrations were bleached out and therefore omitted from the quantitate analysis. The expression time in the BYL system was 16–20 hours at 25 °C, while hCNTF-*Npu*DnaE<sub>ΔC16</sub> was expressed for 6 hours at 30 °C in the HeLa system. Overall, the amounts of hCNTF were 75% and 51% lower than the expected expression amounts of the positive controls of the BYL<sup>17</sup> and the HeLa systems respectively (as given by ThermoFisher Scientific, USA). Please note that due to the absence of β-mercaptoethanol in the loading buffer for the SDS-PAGE, previously purified hCNTF appears as 2 bands on the WB. The major band is the monomeric and the minor band the dimeric form. In previous work Ikonen *et al.* in our laboratory have optimized the expression and purification of hCNTF<sup>18</sup>. Here we can observe one band, the monomeric form of purified hCNTF, in a SDS-PAGE gel with β-mercaptoethanol in the Laemmli buffer, stained with Coomassie brilliant blue.

**Capture.** The capture efficiency onto the beads in the CFPS matrixes was evaluated from the WBs before and after capture with peptides immobilized to magnetic beads from the relative amounts of hCNTF synthesized and remnant in the CFPS matrix (Table 1). The amounts of hCNTF bound to the beads were 0.42 μg/mg beads and

CFPS system	Total amount hCNTF [mg/L] <sup>a</sup>	Amount hCNTF Captured beads [%]	Amount hCNTF Released <sup>b</sup> [%]
BYL <sup>c</sup>	20.0	84	>6.3 <sup>c</sup>
HeLa <sup>c</sup>	49.2	93	ND
HeLa <sup>d</sup>	ND	ND	>19 <sup>c</sup>

**Table 1.** Expression, capture, and light release efficiency of hCNTF with immobilization on the surface of magnetic beads. <sup>a</sup>Milligram hCNTF per liter of CFPS matrix determined with hCNTF standard in the WBs; <sup>b</sup>The percentages are based on the relative volumes of the bands of hCNTF released from the magnetic beads compared to the total synthesized hCNTF in the same WB gels; <sup>c</sup>Immobilized peptide followed by split-intein mediated capture, washing, and light triggered release; <sup>d</sup>split-intein mediated capture and protein *trans* splicing in solution followed by immobilization, washing, and light triggered release; <sup>e</sup>corrected values are based on negative controls (Fig. S4). ND = Not Determined.



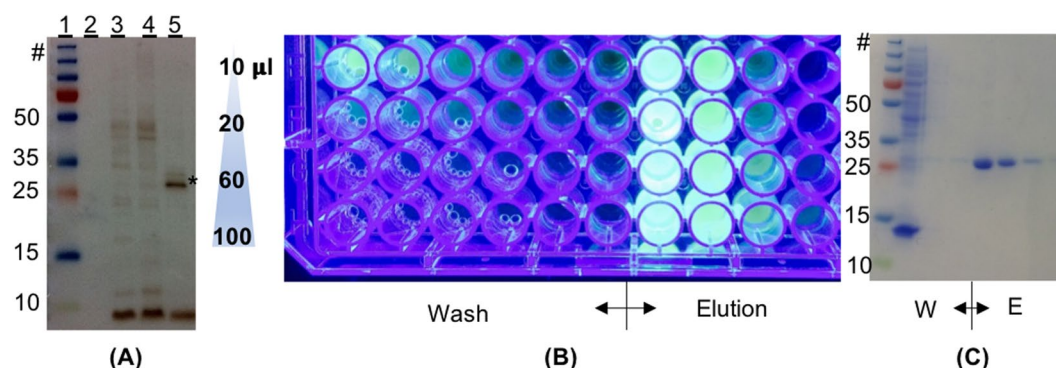
**Figure 2.** Quantitative analysis of the capture and release of hCNTF on the immobilized surface of magnetic beads. **(A)** WB of BYL based CFPS of hCNTF, lanes: (1&15): molecular weight marker; (2–3): total protein before split-intein mediated capture & intein reaction; (4–5): the remnant of the liquid fraction after immobilisation; (6–7): release of hCNTF after 6 hours of light triggered release (365 nm; 12.5 mW); (8–14): 200, 50, 10, 0.5, 0.125 ng of purified His<sub>6</sub>-hCNTF. **(B)** WB of HeLa cell lysate based CFPS of hCNTF, lanes: (1&13): molecular weight marker; (2–3): total protein before split-intein mediated capture & intein reaction; (4–5): the remnant of the liquid fraction after immobilisation; (6–12): 40, 20, 10, 2.5, and 1.3 ng of purified hCNTF. The uncropped images can be seen in Fig. S6.

1.14 µg/mg beads for the BYL and the HeLa matrixes respectively. These values are an order of magnitude below the His<sub>6</sub>-POI binding capacity of the beads of >40 µg GFP/mg beads, indicated by the manufacturer.

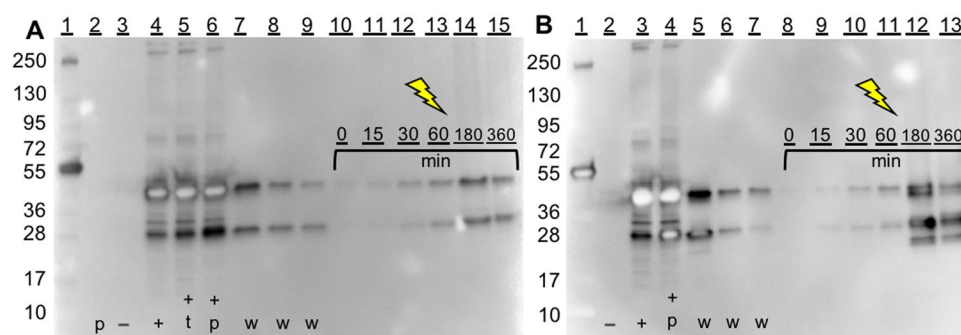
**Purification.** In order to validate split-intein mediated capture and protein *trans* splicing in solution, eGFP-*NpuDnaE*<sub>ΔC16</sub> (enhanced green fluorescent protein-*NpuDnaE*<sub>ΔC16</sub>) was expressed in *E. coli* cells as a recombinant fusion protein. The crude, soluble fraction of the cell lysate was split and either incubated with peptide 1 or peptide 3 to test the feasibility of capture and release in a complex matrix with high protein background, in this case cell lysate. Western blotting was utilized to visualize the target protein with anti-polyhistidine antibody (Fig. 3A). The products eGFP-His<sub>6</sub> (hCNTF-(histidine)<sub>6</sub>) and eGFP-λ-His<sub>6</sub> (hCNTF- photocleavable linker-(histidine)<sub>6</sub>) were purified using affinity chromatography (Fig. 3B,C). The final product eGFP, prepared without a His<sub>6</sub> fusion tag, was clearly purified using affinity chromatography, demonstrating the feasibility of split-intein mediated capture and protein *trans* splicing in a complex matrix. The whole procedure, excluding the cloning experiments, was performed in 6 days.

The washing efficiency of captured and immobilized hCNTF from the Ni-NTA magnetic beads was determined in two scenarios: (a) split-intein mediated capture was performed in the BYL matrix follow by immobilization to Ni-NTA coated magnetic beads, and (b) split-intein mediated capture and protein *trans* splicing in the HeLa lysate matrix with prior immobilized peptides. Vigorous washing of the beads as instructed by the manufacture using the KingFisher (ThermoFisher Scientific) was equally effective for both scenarios and removed (50 ± 18) % of hCNTF from each previous washing step. Thus, before light triggered release, less than 1.5% of the remnant hCNTF was present. Vigorous washing with equal volumes are visualized in Fig. S4A (split-intein mediated capture and protein *trans* splicing with immobilized peptide in HeLa) and S4B (mock-split-intein mediated capture and mock-protein *trans* splicing in BYL). The washing steps depicted in Fig. 4 were less vigorous and concentrated before application to the SDS-PAGE gels in order to visualize if hCNTF was present in the washing step. After washing, hCNTF was cleaved off the beads via light triggered release (Fig. 4A,B) into the final buffer.

**Bioconjugation.** To demonstrate proof-of-principle of moiety transfer from the capture peptide to hCNTF, we utilized peptides 5–7 where biotin was conjugated to the peptides prior to split-intein mediated capture and protein *trans* splicing. Firstly, we transferred biotin to hCNTF synthesized in cell free matrix with peptides 5–7 within the same matrix and observed a migration-shift in the SDS-PAGE gels (Fig. 5A) where the hCNTF-biotin appears to



**Figure 3.** Expression, protein *trans* splicing, and purification of eGFP- *NpuDnaE*<sub>ΔC16</sub> in *E. coli* prior to light triggered release. (A) WB using anti-polyhistidine antibody for the detection in the soluble fraction of *E. coli* lysate. Lanes: (1) molecular weight markers; (2) eGFP- *NpuDnaE*<sub>ΔC16</sub> (not induced: negative control); (3) eGFP- *NpuDnaE*<sub>ΔC16</sub> + peptide 1; (4) eGFP- *NpuDnaE*<sub>ΔC16</sub> + peptide 3 + 1 mM ZnCl<sub>2</sub>; (5) eGFP- *NpuDnaE*<sub>ΔC16</sub> + peptide 3 + 1 mM TCEP. (\*) indicate the mass of eGFP-His6. (B) Elution of the soluble protein fractions after Ni-NTA affinity chromatography, using UV-light for visualization, of the products of the protein *trans* splicing of eGFP- *NpuDnaE*<sub>ΔC16</sub> peptide 1. (C) SDS-PAGE stained with CBB of the first four washing steps and the elution steps. The uncropped images can be seen in Fig. S7.



**Figure 4.** Removal of HeLa cell lysate based CFPS media from the solid support after expression and capture. (A) Split-intein mediated capture (split-intein mediated capture) after CFPS in reaction matrix before immobilization, followed by immobilization, washing, and light triggered release (light triggered release). Lanes: (1) molecular weight marker; (2) Peptide 3; (3) HeLa lysate without DNA (negative control); (4) hCNTF-*NpuDnaE*<sub>ΔC16</sub> CFPS expression without TCEP; (5) hCNTF-*NpuDnaE*<sub>ΔC16</sub> CFPS expression with 1 mM TCEP; (6) hCNTF- *NpuDnaE*<sub>ΔC16</sub> after protein *trans* splicing with peptide 3; (7–9) wash steps 1, 2 and 3; (10–15) 0, 15, 30, 60, 180, and 360 minutes of light triggered release (365 nm, 12.5 mW). (B) Peptide immobilization followed by split-intein mediated capture, washing, and light triggered release. Lanes: (1) molecular weight marker; (2) HeLa lysate without DNA (negative control); (3) hCNTF-*NpuDnaE*<sub>ΔC16</sub> CFPS expression without TCEP; (4) hCNTF- *NpuDnaE*<sub>ΔC16</sub> CFPS expression after protein *trans* splicing with peptide 3; (5–7) wash steps 1, 2 and 3; (8–13) 0, 15, 30, 60, 180, and 360 minutes of light triggered release (365 nm, 12.5 mW). Uncropped images can be seen in Fig. S8.

be smaller, however, it migrates faster due its increased charge. This effect has been described earlier by Sano and Cantor (1990)<sup>19</sup>, where titration of streptavidin with biotin marked an increase in mobility on polyacrylamide gels.

In the streptavidin-shift assay, hCNTF-biotin was detected, though the relative amount appears to be much lower than the difference of hCNTF before and after incubation with streptavidin (Fig. 5B). In Fig. 5B we observe once again the higher mobility of hCNTF-biotin in lane 4, and possibly in lane 6.

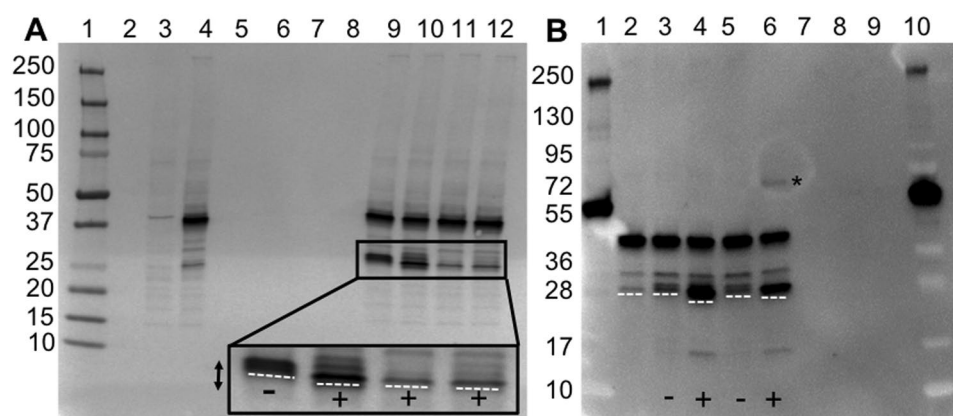
After background correction, 18% of hCNTF is liberated after the split-intein mediated capture and protein *trans* splicing, relative to the amount of hCNTF- *NpuDnaE*<sub>ΔC16</sub>, with peptide 1 (Fig. 5B; lane 3), while peptide 5 is more efficient with 76% liberation of hCNTF. After incubation with streptavidin the relative amount of hCNTF in lane 5 did not change significantly (16%), while the amount of hCNTF-biotin in lane 6 compared to lane 4 is reduced by 56%. In comparison the amount of hCNTF-biotin-streptavidin, marked with an asterisk in lane 6, compared to the amount of hCNTF in lane 4 is 15%.

**Light triggered release.** In order to evaluate the cleavage efficiency of the capture peptide's photocleavable-linkers, the reduced peptides were incubated up to 30 minutes under UV light (365 nm; 12. mW) and analyzed with HPLC (Table 2; Fig. S5). The main peak of the chromatogram after 0 min of UV-light exposure represents 100% of



Peptide number	$t_r$ of main peptide peak (min)	Cleavage efficiency (% of main peptide peak AUC left <sup>a</sup> ) after UV-light exposure (365 nm; 12.5 mW) for:			
		0 min	5 min	15 min	30 min
1	9.975	100	~100	~100	~100
2	10.75	100	72	58	39
3	10.61	100	26	16	8
4	11.84	100	22	11	9
5	10.25	100	~100	~100	~100
6	11.10	100	41	38	30
7	11.80	100	43	39	25

**Table 2.** Efficiency of light cleavage of peptides 1–3 and 5–7 determined with HPLC. <sup>a</sup>See also Fig. S5 for peptides 1–3.



**Figure 5.** Transfer of biotin to hCNTF upon PTS. **(A)** Comparison of peptide 1 and peptides 5–7 (biotin-conjugated peptides). Lanes: (1) molecular weight marker; (2) Laemmli buffer; (3) HeLa lysate without DNA (negative control); (4) HeLa CFPS expressed hCNTF-*NpuDnaE*<sub>ΔC16</sub>; (5) Peptide 1; (6–8) Peptide 5–7; (9–12): hCNTF-*NpuDnaE*<sub>ΔC16</sub> protein *trans* splicing with peptides 1, 5, 6, and 7 respectively. *Inset*: magnification of the hCNTF (–) and hCNTF-biotin bands (+). **(B)** Lanes: (1&10) molecular weight marker; (2) HeLa CFPS expressed hCNTF-*NpuDnaE*<sub>ΔC16</sub> (total protein); (3) HeLa CFPS expressed hCNTF-*NpuDnaE*<sub>ΔC16</sub> with peptide 1 (protein *trans* splicing reaction); (4) HeLa CFPS expressed hCNTF-*NpuDnaE*<sub>ΔC16</sub> with peptide 5 (protein *trans* splicing reaction); (5) HeLa CFPS expressed hCNTF-*NpuDnaE*<sub>ΔC16</sub> with peptide 1 (protein *trans* splicing reaction) and streptavidin; (6) HeLa CFPS expressed hCNTF-*NpuDnaE*<sub>ΔC16</sub> with peptide 5 (protein *trans* splicing reaction) and streptavidin; (7–9) Laemmli buffer; (\*) Streptavidin bound to biotinylated hCNTF. The white dashed lines indicate the front of the bands. Uncropped images can be seen in Fig. S9.

non-cleaved peptide and its integrated peak area was compared to peaks at the same retention time of the same peptide. Peptide 1 and 5 do not contain a photocleavable moiety within the peptide structure, thus any observable decrease in the main peptide peak, and appearance of reaction product's peaks in the chromatogram are solely due to the photocleavable moieties. The most efficient photocleavable group evaluated in this study is 3-(amino)-3-(2-nitrophenyl)propanoic acid (ANP)<sup>20</sup>, while 4-[4-[1-(9-fluorenylmethoxycarbonylamino)ethyl]-2-methoxy-5-nitrophenoxy]butanoic acid (CAS 162827-98-7)<sup>21</sup> is slightly more efficient than 2-nitrophenylalanine (F(2-NO<sub>2</sub>))<sup>22</sup>.

We evaluated the photocleavage after (i) CFPS, (ii) split-intein mediated capture, and (iii) protein *trans* splicing in CFPS matrix followed by washing. Immobilization of hCNTF-λ-His<sub>6</sub> on magnetic beads was also evaluated. hCNTF-*NpuDnaE*<sub>ΔC16</sub> alone was not affected by light treatment (Fig. S4A), while light treatment of the hCNTF-λ-His<sub>6</sub> liberated hCNTF (Figs 2A and 4). The release efficiency presented in Table 1 are conservative percentages. Due to technical limitations, the evaporation of the liquid from the mock-light triggered release samples (Fig. S4A), could not be compensated for, thus we overestimated the amount of hCNTF in the last lane of Fig. S4A of the negative control, resulting in a lower estimate of release of hCNTF in Figs 2A and 4. Hence, we can conclude that the release rates of hCNTF from the magnetic beads (5 mg beads/ml) is at least 6.3% if the peptide 3 is immobilized before split-intein mediated capture and protein *trans* splicing, while the release rate is at least 19% if immobilization occurs after split-intein mediated capture and protein *trans* splicing.

## Discussion

In summary, we demonstrated proof-of-principle of a fast work-flow to prepare, purify and bioconjugate a pharmaceutical relevant hCNTF protein. Cell free expression of proteins is rapid with yields high enough if followed by sensitive screening methods. Here we evaluated a plant-based system tobacco plant BYL and a mammalian-based system based on HeLa cells. The higher yield in the HeLa CFPS system (49 mg/L of hCNTF), compared to BYL was unexpected, but this may be due to the stability of hCNTF. The expression in the HeLa system was done in 6 hours, while the recommended time of 16–20 hours was applied for the BYL system<sup>23</sup>. Initially storage in a 100 mM NaH<sub>2</sub>PO<sub>4</sub>, 50 mM NaCl, 2 mM DTT buffer (pH 8.0) following expression of hCNTF in *E. coli*<sup>18</sup> was not optimal, therefore a ThermoFluor assay<sup>24</sup> was performed. From the thermo-ramping profiles, it was evident that hCNTF unfolds at 42–48 °C (data not shown). For comparison, hCNTF expressed in *E. coli* at a level of >112 mg/L of total protein using a codon optimized gene<sup>18</sup>. The *E. coli* expression and purification was done in 5–6 days.

The capture efficiency of hCNTF-*NpuDnaE*<sub>Δ16</sub> by the capture peptides within the CFPS matrix differed for both systems, mainly due to non-specific binding of proteins already present. This is supported by the negative control (Fig. S4B) where we observe native protein binding and elution to peptide coated Ni-NTA, a known issue when using Ni-NTA surfaces<sup>25</sup>. The capture efficiencies were determined by comparing the relative amounts of hCNTF before and after capture. Split-intein mediated capture onto magnetic beads in CFPS matrix was highly effective up to 93%. The presence of other proteins, therefore, does not seem to limit the capture efficiency. The capture efficiency of the split intein used, in buffer, extrapolated from Fig. S15 from Shah *et al.*<sup>26</sup> appears to be 92–95%. In the same paper, the authors reported that the rate of the intein capture is lower at high salt concentrations. When compared to the specific binding capacities of His<sub>6</sub>-binding resins on magnetic beads, hCNTF is captured with both CFPS systems at a high to very high efficiency at similar bead concentrations<sup>27</sup>. However, the protein amounts bound via split-intein mediated capture followed by protein *trans* splicing are much lower than the capacity of the beads. In the BYL system we immobilized 0.42 μg hCNTF per mg of beads and for the HeLa system this was 1.14 μg hCNTF per mg of beads, while the binding capacity of the beads, according to the manufacturer is over 40 μg GFP per mg of beads. Since our relative binding efficiency is high, but the total amount of hCNTF is relatively low, it seems the maximal capacity of the magnetic beads functionalized with the capture-peptides was not reached.

In the EnBase<sup>TM</sup> cultivation *E. coli* follows a linear growth curve<sup>28</sup>, the total protein background in the *E. coli* lysate was in a similar range of 75–100 fg protein per cell previously reported for slow dividing cells<sup>29</sup>. As shown in Fig. 3, eGFP-*NpuDnaE*<sub>ΔC16</sub> treated with peptide 1 and 3, resulting in eGFP-His<sub>6</sub> and eGFP-λ-His<sub>6</sub>, could be purified by use of Ni-NTA affinity chromatography. Therefore, we demonstrated that split-intein mediated capture and protein *trans* splicing in a complex and high protein background is possible in the crude, soluble fraction of *E. coli* lysate.

Following the capture and binding of hCNTF to Ni-NTA magnetic beads, 5 simple washing steps removed 98.5% of the remaining 7% of hCNTF. After demonstrating the photocleavability of the capture peptides (Table 2) with comparable values reported earlier<sup>20–22</sup>, we demonstrated their functionality in CFPS matrixes. Light triggered release was most effective when protein *trans* splicing was allowed in the complex CFPS matrix before immobilizing to magnetic beads where 19% of immobilized hCNTF was released from the beads due to light treatment. Compared to His<sub>6</sub> fusion tags, coupled with Ni-NTA resin on magnetic beads, their elution (or release) efficiency is between 5–94% for 4 mg beads/ml<sup>27</sup>. Most likely an overcrowded peptide surface is less optimal for a functionally folded intein to form after split-intein mediated capture. During validation experiments of an earlier study (data not show) we found that for flat chitin-SPR sensors approximately 20% of chitin binding domain fused to *NpuDnaE*<sub>ΔC16</sub> was an optimal distribution of the capture peptide<sup>30</sup>. One other limitation of the use of magnetic beads for light triggered release, is the difficulty to penetrate the magnetic bead slurry in the plastic wells of the KingFisher strips (ThermoFisher Scientific, USA) from above with UV-light. Furthermore, though the beads were mixed in a microtiter plate shaker throughout the UV-light exposure in a closed-box, the beads did sediment.

In addition to split-intein mediated capture, protein *trans* splicing, and light triggered release as an alternative protein purification protocol, we also wanted to demonstrate an alternative post-translational modification via intein mediated moiety transfer as an integrated part of the workflow, because bioconjugation is an important aspect of (biological) drug development<sup>31–33</sup>. As shown in the cartoon in Fig. 1, after protein *trans* splicing the POI is covalently linked to the capture peptide. We demonstrated that biotin was transferred from the capture peptide onto hCNTF in one step. In this reaction, no byproducts of the conjugation reaction needed to be removed prior to bioconjugation to the protein<sup>34</sup>. Furthermore, it appears that most of the hCNTF has been bioconjugated (Fig. 5A). Peptide 5 was utilized in a traditional streptavidin-shift assay<sup>35</sup>, since its split-intein mediated capture and protein *trans* splicing was more effective than with peptides 6 and 7 (Fig. 5B). In this reaction approximately 56% of hCNTF was biotinylated, however only 15% of that amount bound to streptavidin. The main reason for this low percentage is the over 100-fold excess of peptide 5 in the reaction mixture. This ‘captures’ a significant amount of the streptavidin away, which in turn cannot bind to the biotinylated hCNTF. In addition, the human anti-hCNTF antibody used for visualization might be less specific for the hCNTF-biotin-streptavidin complex than for hCNTF and hCNTF-biotin respectively. However, even though optimization is needed, moiety transfer after split-intein mediated capture and protein *trans* splicing is clearly demonstrated with cell-free protein synthesized hCNTF and in CFPS matrix.

The current limitations of the workflow depend on the ability to express the POI-intein fusion as a soluble protein and the efficiency of light cleavage while the captured protein resides on magnetic beads. Another current limitation is the yield in absolute amounts. CFPS is scalable<sup>16</sup> and the amount captured is related to the total surface areas covered with the capture peptides. Even if protein production within a CFPS matrix is slower than expected, the CFPS reaction can be prolonged to capture the same amount over a longer period of time.

Currently we are evaluating this new work-flow on flat surfaces to optimize the light triggered release efficiency. We will couple this work-flow with process analytical techniques to evaluate the amount and quality of the pharmaceutical protein produced as well as in-line assays relevant to its function. Finally, we will indent to scale up the capture surface to increase the yields.

## Methods

**Materials.** Unless specifically mentioned all commercial chemicals tris(2-carboxyethyl)phosphine (TCEP), ethylenediaminetetraacetic acid (EDTA), dithiothreitol (DTT), sodium azide ( $\text{NaN}_3$ ), triethanolamine hydrochloride (TEA), tris, sodium chloride, ampicillin, isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), magnesium chloride, imidazole, 2-(N-morpholino)ethanesulfonic acid (MES), and lysozyme were obtained from Merck Sigma-Aldrich (Darmstadt, Germany), Potassium chloride was obtained from Honeywell Riedel de Haën (Seelz, Germany), disodium hydrogen phosphate was obtained from Fisher Scientific (Hampton, USA). DNase and RNase were obtained from Epicentre (Madison, WI, USA). Phenylmethylsulfonyl fluoride (PMSF) was obtained from ThermoFisher Scientific, USA. Peptides were synthesized by Peptide Protein Research Ltd, Hampshire, U.K. Following reaction buffers and media were used: Buffer A: 100 mM MES, 500 mM NaCl, pH 7.0 (at 25 °C); Buffer B: 50 mM Tris/HCl, 50 mM NaCl, pH 8.0 (at 25 °C); Buffer C: 10 mM Tris/HCl, 150 mM NaCl, 1 mM TCEP, pH 7.4 (at 25 °C). EnPressoB medium was obtained from BioSilta Ltd. (U.K); currently available via EnPresso GmbH (Berlin, DE). Ultra-pure water, having a resistivity of 18 M $\Omega$ -cm and a total organic carbon level of <5 ppm, prepared with a Milli-Q purification system was used for all measurements and for preparation of the buffers, the peptide solutions, and EnPressoB media.

**Construction of expression vectors.** Yeast codon optimized human heat shock protein A1 (HspA1) was purchased from ATUM<sup>SM</sup> (then DNA2.0, USA) and the HspA1 gene was cloned into the pIVEX1.4 Strep tag II – Tobacco Etch virus protease (TEV) plasmid by Biotechnrabbt GmbH (DE). The pIVEX\_GFP plasmid was purchased from Biotechnrabbt GmbH (DE) as component of an *E. coli* CFPS expression kit (order number BR1400101).

The codon optimized hCNTF gene was isolated from pOPINF-CNTF plasmid (Itkonen *et al.* 2014) in a gradient-PCR (57–62 °C) with 5'-ttaaagaaggagatatacatATGGCGTTTACCGAACATTC-3' and 5'-gtttcatagctgaggcaCATCTTCTTGTGTTTGGCG-3' as forward and reverse primers, while the eGFP gene was isolated from pEGFP-C1 plasmid (Clontech, USA) in a PCR with 5'- ttaaagaaggagatatacatATGGTGAGCAAGGGCGAG-3' and 5'-gtttcatagctgaggcaCTGTACAGCTCGTCCATGC-3' as forward and reverse primers. The NpuDnaE<sub>NA16</sub> harbouring pET21a-CBD-NpuDnaE<sub>NA16</sub> plasmid<sup>30</sup> was linearized using 5'-TGCCTCAGCTATGAAACG-3' and 5'-ATGTATATCTCCTTCTTAAAGTTAAAC-3' as forward and reverse primers. The PCR products were assembled into plasmids pET21a-CNTF-NpuDnaE<sub>NA16</sub> and pET21a-eGFP-NpuDnaE<sub>NA16</sub> with NEBuilder<sup>®</sup>.

Plasmids pT7CFE1-CHis (Thermo Scientific, Rockford, IL, USA) and pIVEX\_GAA\_Omega\_eYFP\_His<sup>17,23</sup> were used as vector backbones in constructing the expression plasmids for the CFPS of hCNTF-NpuDnaE<sub>CA16</sub> in the HeLa and Tobacco BYL CFPS systems, respectively. DNA fragments for the CFPS expression plasmids were generated as follows. The hCNTF-NpuDnaE<sub>CA16</sub> sequence from the pET21-hCNTF-NpuDnaE<sub>CA16</sub> plasmid was amplified using gradient-PCR (57–62 °C) with 5'-gatgataatgcccaccaccatATGGCGTTTACCGAAC-3' and 5'-ttttttt-ttttttttttCTAGCGCTCAACTCCAATGTC-3' (Fragment A), and 5'-acattttacatttacaactacc-ATGGCGTTTACCGAACATTC-3' and 5'-cagaagagctgggtacctaTTAGCGCTCAACTCCAATG-3' (Fragment B) as forward and reverse primers as for the HeLa and BYL constructs, respectively. The sequence harboring the C-terminal His<sub>6</sub>-tag was cloned out and the pT7CFE1-CHis plasmid linearized by restriction enzyme digestion with MscI and BglII (both NEB, MA, USA) (Fragment C). The sequence for eYFP-His was removed and the linearized pIVEX\_GAA\_Omega\_ plasmid backbone amplified in an inverse gradient-PCR (64–69 °C) with 5'-TAAGGTACCCAGCTCTTC-3' and 5'-GGTAGTTGTAGAATGTAAAA-TGTAATG-3' as forward and reverse primers, respectively (Fragment D). Fragments A + C and B + D were assembled into plasmids pT7CFE1-hCNTF-NpuDnaE<sub>CA16</sub> and pIVEX\_GAA\_Omega\_hCNTF-NpuDnaE<sub>CA16</sub>, respectively, using NEBuilder<sup>®</sup> (NEB, MA, USA).

All DNA fragments were separated on 1% agarose gels, and bands of the appropriate sizes excised and cleaned up with NucleoSpin<sup>®</sup> Gel and PCR Clean-up columns (Macherey-Nagel, DE). The purified DNA fragments were assembled into plasmids using NEBuilder<sup>®</sup> HiFi DNA Assembly Cloning Kit (NEB, MA, USA) and the yielded expression plasmids were transformed into NEB-5 $\alpha$  (NEB, MA, USA) *E. coli* competent cells<sup>36</sup>. Multiple colonies were picked and screened for the insertion of the hCNTF-NpuDnaE<sub>CA16</sub> sequence by colony PCR using 5'-TAATACGACTCACTATAGGG-3' and 5'-tttttttttttttttttttCTAGCGCTCAACTCCAATGTC-3', and 5'-ATGTAATACGACTCACTATAGAAA-3' and 5'-AGGTCCAAACCAACCA-3' as forward and reverse primers for HeLa and BYL constructs, respectively. Positive clones were propagated and the sequences of the amplified plasmids verified by gene sequencing (GATC Biotech, DE). Verified plasmid constructs were further amplified and purified to high concentration and purity using NucleoBond<sup>®</sup> Xtra Midi columns (Macherey-Nagel, DE). Plasmids were then transformed separately to *E. coli* strain BL21(DE3) (Invitrogen, USA) for protein production and in parallel stored at –20 °C for CFPS.

**Cellular protein expression and purification.** Expression and purification of His<sub>6</sub>-hCNTF used for the controls was performed as previously described<sup>18</sup>. Briefly summarized, His<sub>6</sub>-hCNTF was produced in Rosetta 2 (DE3) pLysS with the auto-inducing media TBONEX (450 ml in 2000 ml baffled flask; 25 °C at 270 rpm for 30 h). Frozen cell pellets were freeze thawed and resuspended (3x) in 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 300 mM NaCl, 10 mM imidazole, 1 mM DTT, protease inhibitor cocktail, and lysonase (pH 8.0). His<sub>6</sub>-hCNTF present in the soluble cell lysate fraction was purified with a Ni-IDA resin and size exclusion chromatography. The protocol had one adaption: a ThermoFluor assay<sup>24</sup> was performed to find a buffer (buffer A) which provided a better His<sub>6</sub>-hCNTF stability



Peptide number	<i>NpuDnaE</i> <sub>ΔC16</sub> peptide sequence with non-canonical amino acid/ conjugation in bold	Non-canonical amino acid/ conjugation	Peptide sequence introduced on C-terminus of POI via PTS
1	DGHNFKNGFIASNCFGSKHHHHHH	N/A	CFGSKHHHHHH
2	DGHNFKNGFIASNC <b>F(2-NO<sub>2</sub>)</b> GSKHHHHHH	F(2-NO <sub>2</sub> ): 2-Nitrophenylalanine <sup>22</sup>	<b>CF(2-NO<sub>2</sub>)</b> GSKHHHHHH
3	DGHNFKNGFIASNC- <b>ANP</b> -GSKHHHHHH	ANP: 3-(Amino)-3-(2-nitrophenyl) propionic acid <sup>20</sup>	<b>C-ANP</b> -GSKHHHHHH
4	DGHNFKNGFIASNC <b>F-X</b> -GSKHHHHHH	X: Photolinker CAS 162827-98-7 <sup>21</sup>	<b>CF-X</b> -GSKHHHHHH
5	DGHNFKNGFIASNC <b>GSX</b> -GSKHHHHHH	X: Biotin on side chain of Lysine	<b>CGSG-X</b> -GSKHHHHHH
6	DGHNFKNGFIASNC <b>GSX-XY</b> -GSKHHHHHH	X: Biotin on side chain of Lysine Y: Photolinker CAS 162827-98-7 <sup>21</sup>	<b>CGSG-XY</b> -GSKHHHHHH
7	DGHNFKNGFIASNC <b>GSX-XY</b> -GGEPEA	X: Biotin on side chain of Lysine Y: Photolinker CAS 162827-98-7	<b>CGSG-XY</b> -GGEPEA

**Table 3.** Photocleavable peptides used in this study.

than the original storage buffer. Expression of eGFP-His6-*NpuDnaE*<sub>ΔC16</sub> was carried out in BL21(DE3) *E. coli* cells. First the transformants were grown o/n in LB Medium with 100 µg/ml ampicillin and 1% (w/v) glucose at 30 °C from glycerol stocks; 2 ml o/n culture was used to inoculate 50 ml of EnPressoB medium<sup>28</sup> with 100 µg/ml ampicillin at 30 °C; 225 rpm (1" amplitude shaker) according to manufacturer's instructions in high yield flasks<sup>37</sup>. Cells were induced with 0.4 mM IPTG after 24 h and grown for another 24 h under the same conditions. The cells were separated from the media with centrifugation (16 000 \* g for 15 min) and then lysed by freezing o/n at -20 °C in the presence of DNase (2.5 U/ml), RNase (2.5 U/ml), lysozyme (15 µg/ml), 0.1 mM PMSF (ThermoFisher Scientific, USA), and MgCl<sub>2</sub> in buffer A. The insoluble fraction was removed by centrifugation (16000 g for 45 min). The crude, soluble *E. coli* lysate was kept on ice until further use. After incubating the cell lysates with either peptide 1 or peptide 3 Table 3) for 3 hours, the products of protein *trans* splicing were purified with a Protino® 96 Ni-NTA well plate with using the vacuum manifold (Macherey-Nagel, Germany) according to the manufactures instructions using buffer A and recommended imidazole concentrations. Different wash volumes were evaluated and performed (N = 6), elution was as instructed and was performed four-fold. The resulting protein fractions of the wash and eluate steps were visualized with SDS-PAGE gels using coomassie brilliant blue R-250 (ThermoFisher Scientific, USA)

**Cell free protein synthesis (CFPS).** *E. coli* BL21(DE3) S12 and S30 lysates<sup>38–41</sup>, tobacco plant or BYL<sup>17</sup>, *E. coli* (Promega, USA, order number L1110), wheat germ (Biotech rabbit, Germany, order number BR1401001), and mammalian HeLa cells (ThermoFisher Scientific, USA, order number 88881) were utilized for protein expression according to the published protocols or manufacturer's instructions. The quality of each CFPS reaction was checked by use of a positive control protein (GFP) in parallel to the CFPS reactions of the samples. Expression levels were also compared to a negative control where no DNA was added to the reaction. Continuous-exchange cell-free expression (CECF) of Strep-Tag II \_TEV- HspA1 was performed in wheat germ lysates. CFPS of GFP was performed in *E. coli* BL21(DE3) S12 and S30 lysates with the PremixPlus reaction mixture from the Promega *E. coli* kit and *E. coli* lysate (Promega, USA) with the PremixPlus, the 'Swartz' reaction mix<sup>41</sup>, and the 'EMBL' reaction mix<sup>39</sup>. CFPS of hCNTF-*NpuDnaE*<sub>ΔC16</sub> was carried out in the BYL and HeLa systems. Briefly, required reaction components were mixed, plasmid DNA template was added and the reaction mixtures incubated accordingly; BYL CFPS reactions were incubated in an incubator-shaker (25 °C for 20 h) in 96-well plates while HeLa CFPS reactions took place at 30 °C during 6 h in 1.5 ml microcentrifuge tubes. The synthesized proteins in the CFPS matrix were used in subsequent experiments immediately.

**Capture and Release.** The validity of the intein-mediated protein *trans* splicing of the artificially split *NpuDnaE* intein (*NpuDnaE*<sub>ΔC15</sub>/*NpuDnaE*<sub>ΔC15</sub><sup>42,43</sup>; see availability of data and materials) in the capture and release of expressed protein was studied. The expressed, captured and released hCNTF was studied with Western Blotting. To verify that protein *trans* splicing can indeed ligate a C-terminal peptide sequence with e.g. a photocleavable moiety and a His<sub>6</sub>-tag (for the capture and release) to the expressed POI, the following experiments were set up. PTS reactions between photocleavable peptide 2/peptide 3 (Table 3) and CFPS reactions containing expressed hCNTF-*NpuDnaE*<sub>ΔC16</sub> were carried out for 3 hours at room temperature. The resulting putative His<sub>6</sub>-tagged hCNTF was captured with a 1 mg of HisPur Ni-NTA Magnetic Beads, corresponding to a 5–10 mg/ml final concentrations of beads (ThermoFisher Scientific, USA). All handling of the magnetic beads was carried out with the KingFisher Purification System (ThermoFisher Scientific, USA). The unbound proteins and components of the CFPS reactions were subsequently removed from the beads by five washing-steps with 100 µl 0.05% Tween-20 in PBS (pH 8.0). Release of hCNTF was carried out by photocleavage with UV light (365 nm; 12.5 mW) for 0–360 min.

The feasibility of utilizing the intein ligation reaction itself in the capture and subsequent release of the expressed protein was also studied. In this setup, peptide 3 was first immobilized on the Ni-NTA beads. Intein ligation reaction between immobilized peptide 3 and CFPS expressed hCNTF-*NpuDnaE*<sub>ΔC16</sub> was carried out overnight in room temperature, resulting in the ligation product, hCNTF-His<sub>6</sub>, that remains immobilized on the magnetic beads. After removal of unbound proteins and CFPS reaction components by wash steps, release of hCNTF was carried out by photocleavage with UV light (365 nm; 12 mW) for 0–360 min.

To verify that the capture and release were indeed achieved *via* the intein ligation reaction, control experiments were carried out. To rule out nonspecific binding to the Ni-NTA beads, a mock binding experiment was carried out using only the CFPS expressed hCNTF- *NpuDnaE*<sub>ΔC16</sub> with the beads. After removal of unbound proteins and components of the CFPS reaction by wash steps, 'release' was carried out by photocleavage with UV light (365 nm; 12.5 mW) for 0–360 min. To verify the release having taken place *via* cleavage of the photolabile unnatural amino acids, the non-photocleavable peptide 1 was first immobilized on the Ni-NTA beads. Intein ligation reaction between immobilized peptide 1 and CFPS expressed hCNTF- *NpuDnaE*<sub>ΔC16</sub> was carried out overnight in room temperature, resulting in the ligation product, hCNTF-His<sub>6</sub>, staying immobilized on the magnetic beads. After removal of unbound proteins and components of the CFPS reaction by wash steps, 'release' of hCNTF was carried out by photocleavage with UV light (365 nm; 12.5 mW) for 0–360 min.

**Bioconjugation.** Protein *trans* splicing reactions were carried out between expressed hCNTF- *NpuDnaE*<sub>ΔC16</sub> and *NpuDnaE*<sub>ΔC16</sub> peptide 1 or modified peptides 5, 6 or 7; the biotinylated peptide 5, peptide 6 and peptide 7 were used to assess if the respective modifications (Table 3) affect the Protein *trans* splicing e.g. through steric hindrance, while the non-conjugated peptide 1 was used as a positive control for the reaction. Briefly, hCNTF- *NpuDnaE*<sub>ΔC16</sub> in HeLa CFPS reaction matrix was mixed with peptide 1, peptide 5, peptide 6 or peptide 7 (all 100 μM; reduced with 25 mM TCEP at 27 °C for 30 min) in buffer C and the PTS allowed to take place in room temperature for 2 h. Reactions were halted with the addition of 4X Laemmli sample buffer.

In a similar setup, peptide 5 was used to assess transferring biotin onto the protein of interest *via* PTS, while peptide 2 was used as a negative control. PTS reactions were set up as described earlier and carried out at room temperature overnight. Taking advantage of the well-documented strong and SDS-resistant binding between streptavidin and biotin, a streptavidin gel-shift assay (adapted from Fairhead and Howarth and Sorensen *et al.*<sup>35,44</sup>) was used to confirm transfer of biotin moiety to hCNTF; the described PTS reaction setups yielding the putative non-biotinylated and biotinylated hCNTF products were mixed with streptavidin in PBS and the streptavidin-biotin binding allowed 3 hours in room temperature. Detection of hCNTF biotinylation and subsequent streptavidin binding was carried out by Western Blotting.

**HPLC.** To study the photocleavage efficiency and rate of synthesized *NpuDnaE*<sub>ΔC16</sub>-peptides with varying linkers and conjugates, the peptide samples were dissolved in ultra-pure water and reduced with 25 mM TCEP at 27 °C for 30 min and then exposed to UV-light (365 nm; 12.5 mW) for 0–30 min. The generated peptide traces were separated with an Agilent 1100 HPLC system (Agilent Technologies, Palo Alto, CA) with UV detection at 216 nm. A Discovery BIO Wide Pore C18 (150 × 4.6 mm, 5 μm; Sigma-Aldrich) column maintained at 25 °C and at a flow rate of 1.5 ml/min was used for all analyses. Samples (15 μl) were injected into the column using water (A) and ACN (B), both with 1% TFA, as eluents. The gradient program (0–14 min 1 → 40% B, 14–16 min 40 → 70% B, 16–16.1 min 70 → 1% B, followed by 14 min equilibrium at 1% B) was used to elute all peptide samples. Resulting chromatograms were examined and processed with ChemStation software A.08.03 (Agilent Technologies); decrease in the AUC of the respective main peptide peak was used as a surrogate to measure light triggered cleavage of the different peptides.

**Western blotting.** Protein samples were resolved on precast 4–20% Mini-Protein<sup>®</sup> TGX Stain-Free SDS-PAGE gels (Bio-Rad) with PageRuler<sup>®</sup> Prestained Protein Ladder (ThermoFisher Scientific, USA) or Precision Plus Protein WesternC Standard (Bio-Rad, USA) used as molecular weight markers. Separated proteins were transferred onto 0.2 μm nitrocellulose membranes with the Trans-Blot Turbo Transfer System (Bio-Rad) according to manufacturer's instructions and the membranes blocked in 2% bovine serum albumin (BSA) in Tris-buffered saline, 0.05% Tween 20 (TBS-T) for 2 h in room temperature or overnight in +4 °C. Proteins were detected in a sandwich reaction with 0.1–0.2 μg/ml rabbit anti-hCNTF polyclonal antibody (ThermoFisher Scientific, USA) and 0.1 μg/ml horseradish peroxidase-conjugated goat anti-rabbit IgG (Merck Sigma Aldrich, DE) as primary and secondary antibodies, respectively. Staining was carried out with Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare, USA) and the chemiluminescence detected with ChemiDoc XRS+ imaging system (Bio-Rad, USA). The band intensities were quantified using with the Image Lab software (Bio-Rad, USA)<sup>45</sup>. In brief, we image the WB with UV light and subtract the background after inspecting each lane. Known hCNTF quantities were used to calculate the amount of hCNTF on the sample gels. Overexposed band were excluded from the quantitative analysis. Membranes were rinsed with TBS-T between incubations.

**Availability of data and materials.** The datasets supporting the conclusions of this article are included within the article and the supplementary data file.

The intein nomenclature is according to Aranko *et al.*<sup>46</sup>, however since our capture peptides are 16 amino acids long, we use *NpuDnaE*<sub>ΔC16</sub> instead of *NpuDnaE*<sub>ΔC15</sub>, however the split-site is the same.

## References

- Robert, B. *The Uses of Life: A History of Biotechnology*. (Cambridge University Press, 1994).
- Chou, C. P. Engineering cell physiology to enhance recombinant protein production in *Escherichia coli*. *Appl Microbiol Biotechnol* **76**, 521–532, <https://doi.org/10.1007/s00253-007-1039-0> (2007).
- Hunt, I. From gene to protein: a review of new and enabling technologies for multi-parallel protein expression. *Protein Expr Purif* **40**, 1–22, <https://doi.org/10.1016/j.pep.2004.10.018> (2005).
- Palomares L. A., E.-M. S., Ramirez O. T. In *Recombinant Gene Expression. Methods in Molecular Biology* Vol. 267 (ed. Lorence A. Balbás P.) (Humana Press, 2004).
- Structural Genomics, C. *et al.* Protein production and purification. *Nat Methods* **5**, 135–146, <https://doi.org/10.1038/nmeth.f.202> (2008).
- Casteleijn, M. G. & Richardson, D. Engineering Cells and Proteins – creating pharmaceuticals. *Eur. Pharm. Rev.* **19**, 7 (2014).

7. Cariuk, P., Gardener, M. J. & Vaughan, T. J. Evolution of biologics screening technologies. *Pharmaceuticals (Basel)* **6**, 681–688, <https://doi.org/10.3390/ph6050681> (2013).
8. Gronemeyer, P., Ditz, R. & Strube, J. Trends in Upstream and Downstream Process Development for Antibody Manufacturing. *Bioengineering (Basel)* **1**, 188–212, <https://doi.org/10.3390/bioengineering1040188> (2014).
9. Enfors, S. O. & Häggström, L. *Bioprocess technology fundamentals and applications*. (Hörskolettryckeriet, Royal Institute of Technology, 2000).
10. Labrou, N. E. Protein purification: an overview. *Methods Mol Biol* **1129**, 3–10, [https://doi.org/10.1007/978-1-62703-977-2\\_1](https://doi.org/10.1007/978-1-62703-977-2_1) (2014).
11. Burgess, R. R. Protein precipitation techniques. *Methods Enzymol* **463**, 331–342, [https://doi.org/10.1016/S0076-6879\(09\)63020-2](https://doi.org/10.1016/S0076-6879(09)63020-2) (2009).
12. McCue, J. T., Selvitelli, K. & Walker, J. Application of a novel affinity adsorbent for the capture and purification of recombinant factor VIII compounds. *J Chromatogr A* **1216**, 7824–7830, <https://doi.org/10.1016/j.chroma.2009.09.045> (2009).
13. Peti, W. & Page, R. Strategies to maximize heterologous protein expression in *Escherichia coli* with minimal cost. *Protein Expr Purif* **51**, 1–10, <https://doi.org/10.1016/j.pep.2006.06.024> (2007).
14. Landowski, C. P. *et al.* Combinatorial use of antibodies to secreted mycobacterial proteins in a host immune system-independent test for tuberculosis. *J Clin Microbiol* **39**, 2418–2424, <https://doi.org/10.1128/JCM.39.7.2418-2424.2001> (2001).
15. Shields, K. M., Panzhinskiy, E., Burns, N., Zawada, W. M. & Das, M. Mitogen-activated protein kinase phosphatase-1 is a key regulator of hypoxia-induced vascular endothelial growth factor expression and vessel density in lung. *Am J Pathol* **178**, 98–109, <https://doi.org/10.1016/j.ajpath.2010.11.025> (2011).
16. Casteleijn, M. G., Urtti, A. & Sarkhel, S. Expression without boundaries: cell-free protein synthesis in pharmaceutical research. *Int J Pharm* **440**, 39–47, <https://doi.org/10.1016/j.ijpharm.2012.04.005> (2013).
17. Buntru, M., Vogel, S., Stoff, K., Spiegel, H. & Schillberg, S. A Versatile Coupled Cell-Free Transcription-Translation System Based on Tobacco BY-2 Cell Lysates. *Biotechnology and Bioengineering* **112**, 867–878, <https://doi.org/10.1002/bit.25502> (2015).
18. Itkonen, J. M., Urtti, A., Bird, L. E. & Sarkhel, S. Codon optimization and factorial screening for enhanced soluble expression of human ciliary neurotrophic factor in *Escherichia coli*. *BMC Biotechnol* **14**, 92, <https://doi.org/10.1186/s12896-014-0092-x> (2014).
19. Sano, T. & Cantor, C. R. Cooperative Biotin Binding by Streptavidin - Electrophoretic Behavior and Subunit Association of Streptavidin in the Presence of 6-M Urea. *Journal of Biological Chemistry* **265**, 3369–3373 (1990).
20. Brown, B. B., Wagner, D. S. & Geysen, H. M. A single-bead decode strategy using electrospray ionization mass spectrometry and a new photolabile linker: 3-Amino-3-(2-nitrophenyl)propionic acid. *Molecular Diversity* **1**, 4–12, <https://doi.org/10.1007/bf01715804> (1995).
21. Whitehouse, D. L., Savinov, S. N. & Austin, D. J. An improved synthesis and selective coupling of a hydroxy based photolabile linker for solid phase organic synthesis. *Tetrahedron Lett* **38**, 7851–7852, [https://doi.org/10.1016/S0040-4039\(97\)10133-2](https://doi.org/10.1016/S0040-4039(97)10133-2) (1997).
22. England, P. M., Lester, H. A., Davidson, N. & Dougherty, D. A. Site-specific, photochemical proteolysis applied to ion channels *in vivo*. *P Natl Acad Sci USA* **94**, 11025–11030, <https://doi.org/10.1073/pnas.94.20.11025> (1997).
23. Buntru, M., Vogel, S., Spiegel, H. & Schillberg, S. Tobacco BY-2 cell-free lysate: an alternative and highly-productive plant-based *in vitro* translation system. *BMC Biotechnol* **14**, 37, <https://doi.org/10.1186/1472-6750-14-37> (2014).
24. Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N. & Nordlund, P. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* **357**, 289–298, <https://doi.org/10.1016/j.ab.2006.07.027> (2006).
25. Bolanos-Garcia, V. M. & Davies, O. R. Structural analysis and classification of native proteins from *E. coli* commonly co-purified by immobilised metal affinity chromatography. *Biochim Biophys Acta* **1760**, 1304–1313, <https://doi.org/10.1016/j.bbagen.2006.03.027> (2006).
26. Shah, N. H., Eryilmaz, E., Cowburn, D. & Muir, T. W. Naturally split inteins assemble through a “capture and collapse” mechanism. *J Am Chem Soc* **135**, 18673–18681, <https://doi.org/10.1021/ja4104364> (2013).
27. Nejadmoghaddam, M. R., Chamankhah, M., Zarei, S. & Zarnani, A. H. Profiling and quantitative evaluation of three nickel-coated magnetic matrices for purification of recombinant proteins: helpful hints for the optimized nanomagnetisable matrix preparation. *J Nanobiotechnology* **9**, 31, <https://doi.org/10.1186/1477-3155-9-31> (2011).
28. Panula-Perala, J. *et al.* Enzyme controlled glucose auto-delivery for high cell density cultivations in microplates and shake flasks. *Microb Cell Fact* **7**, 31, <https://doi.org/10.1186/1475-2859-7-31> (2008).
29. Wisniewski, J. R. & Rakus, D. Multi-enzyme digestion FASP and the ‘Total Protein Approach’-based absolute quantification of the *Escherichia coli* proteome. *J Proteomics* **109**, 322–331, <https://doi.org/10.1016/j.jprot.2014.07.012> (2014).
30. Casteleijn, M. G. *et al.* Spin coated chitin films for biosensors and its analysis are dependent on chitin-surface interactions. *Colloid Surface A* **539**, 261–272, <https://doi.org/10.1016/j.colsurfa.2017.12.036> (2018).
31. Veronese, F. M. & Morpurgo, M. Bioconjugation in pharmaceutical chemistry. *Farmaco* **54**, 497–516 (1999).
32. Wang, M. *et al.* Amino acids/peptides conjugated heterocycles: A tool for the recent development of novel therapeutic agents. *Bioorg Chem* **76**, 113–129, <https://doi.org/10.1016/j.bioorg.2017.11.007> (2018).
33. Vhora, I., Patil, S., Bhatt, P. & Misra, A. Protein- and Peptide-Drug Conjugates: An Emerging Drug Delivery Technology. *Adv Protein Chem Str* **98**, 1–55, <https://doi.org/10.1016/bs.apcsb.2014.11.001> (2015).
34. Stephanopoulos, N. & Francis, M. B. Choosing an effective protein bioconjugation strategy. *Nat Chem Biol* **7**, 876–884, <https://doi.org/10.1038/Nchembio.720> (2011).
35. Sorenson, A. E., Askin, S. P. & Schaeffer, P. M. In-gel detection of biotin-protein conjugates with a green fluorescent streptavidin probe. *Analytical Methods* **7**, 2087–2092, <https://doi.org/10.1039/C4AY02666G> (2015).
36. Sambrook, J. & Russell, D. W. *Molecular cloning: a laboratory manual*. 3rd edn, (Cold Spring Harbor Laboratory Press, 2001).
37. Ukkonen, K., Vasala, A., Ojamo, H. & Neubauer, P. High-yield production of biologically active recombinant protein in shake flask culture by combination of enzyme-based glucose delivery and increased oxygen transfer. *Microb Cell Fact* **10**, 107, <https://doi.org/10.1186/1475-2859-10-107> (2011).
38. Kigawa, T. *et al.* Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. *J Struct Funct Genomics* **5**, 63–68, <https://doi.org/10.1023/B:JSFG.0000029204.57846.7d> (2004).
39. EMBL. *Protein Expression. E.coli: How to Prepare your Own Cell-Free System*, [https://www.embl.de/pepcore/pepcore\\_services/protein\\_expression/ecoli/lysate/index.html#setting](https://www.embl.de/pepcore/pepcore_services/protein_expression/ecoli/lysate/index.html#setting) (2015 - 2018).
40. Kim, T. W. *et al.* Simple procedures for the construction of a robust and cost-effective cell-free protein synthesis system. *J Biotechnol* **126**, 554–561, <https://doi.org/10.1016/j.jbiotec.2006.05.014> (2006).
41. Jewett, M. C., Calhoun, K. A., Voloshin, A., Wu, J. J. & Swartz, J. R. An integrated cell-free metabolic platform for protein production and synthetic biology. *Mol Syst Biol* **4**, 220, <https://doi.org/10.1038/msb.2008.57> (2008).
42. Aranko, A. S., Zuger, S., Buchinger, E. & Iwai, H. *In Vivo* and *In Vitro* Protein Ligation by Naturally Occurring and Engineered Split DnaE Inteins. *Plos One* **4**, <https://doi.org/10.1371/journal.pone.0005185> (2009).
43. Aranko, A. S. *et al.* Structure-based engineering and comparison of novel split inteins for protein ligation. *Mol Biosyst* **10**, 1023–1034, <https://doi.org/10.1039/c4mb00021h> (2014).
44. Fairhead, M. & Howarth, M. Site-specific biotinylation of purified proteins using BirA. *Methods Mol Biol* **1266**, 171–184, [https://doi.org/10.1007/978-1-4939-2272-7\\_12](https://doi.org/10.1007/978-1-4939-2272-7_12) (2015).
45. Taylor, S. C., Berkelman, T., Yadav, G. & Hammond, M. A defined methodology for reliable quantification of Western blot data. *Mol Biotechnol* **55**, 217–226, <https://doi.org/10.1007/s12033-013-9672-6> (2013).
46. Aranko, A. S., Wlodawer, A. & Iwai, H. Nature’s recipe for splitting inteins. *Protein Eng Des Sel* **27**, 263–271, <https://doi.org/10.1093/protein/gzu028> (2014).
47. Wu, H., Hu, Z. & Liu, X. Q. Protein trans-splicing by a split intein encoded in a split DnaE gene of *Synechocystis* sp. PCC6803. *Proc Natl Acad Sci USA* **95**, 9226–9231 (1998).

## Acknowledgements

We thank Dr. Hideo Iwai of the Institute of Biotechnology (University of Helsinki) for providing us with the plasmid containing the *NpuDnaE* gene, the Institute of Biotechnology Crystallization Facility for assistance with the ThermoFluor analysis, and Dr. Matthias Buntru of the Fraunhofer Institute for Molecular Biology and Applied Ecology (Germany) for his kind contribution of providing us with the tobacco BYL CFPS system. We also would like to thank Leena Pietilä for her kind laboratory assistance, and Regina Casteleijn-Osorno of AaltoEE (Finland) for comments that greatly improved the manuscript. We kindly acknowledge our funding sources and support from the Academy of Finland projects 272597 and 303884. J.I. has also been funded by grants from the Finnish Cultural Foundation, the Evald and Hilda Nissi Foundation, the Päivikki and Sakari Sohlberg Foundation, and the Paulo Foundation. J.N. received Erasmus+ program funding from the Spanish Ministry of Education, Culture & Sports.

## Author Contributions

D.R. prepared genetic constructs, expressed and purified protein, developed the HPLC method, and analyzed and interpreted the capture and release data in *E. coli* lysates and buffer. J.I. prepared genetic constructs, expressed and purified protein, and analyzed and interpreted the capture and release data in CFPS matrix and release profiles from solid surfaces; in addition, he performed the HPLC analysis. J.N. prepared S12 and S30 *E. coli* lysates and compared *E. coli* lysate-based expression levels. A.U. provided funding, feed-back, and supervision. M.G.C. expressed and purified protein, and analyzed and interpreted the capture and release data in *E. coli* lysates, buffer and CFPS, provided funding and was a major contributor in writing the manuscript. All authors contributed to the writing, have read and approved the final manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-30435-4>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018